Homework 5

Exercise 2. Diffusion kernel on a grid

1) • First, let us write L_1 and L_2 as matrices, using $L = D - A$ and the different connections between points in the line graph / square grids

$$L_1 = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & 2 & -1 \\ & & -1 & 1 \end{bmatrix}_{m \times m} \qquad L_2 = \begin{bmatrix} M & -I & & \\ -I & N & \ddots & \\ & \ddots & N & -I \\ & & -I & M \end{bmatrix}_{m^2 \times m^2}$$

where $M = \begin{bmatrix} 2 & -1 & & \\ -1 & 3 & \ddots & \\ & \ddots & 3 & -1 \\ & & -1 & 2 \end{bmatrix}$, $N = \begin{bmatrix} 3 & -1 & & \\ -1 & 4 & \ddots & \\ & \ddots & 4 & -1 \\ & & -1 & 3 \end{bmatrix}$

(we index the points row by row for the grid)

One can remark that $M = I + L_1$ and $N = 2I + L_1$

• Now, let us show that $e_i \otimes e_j$ are the eigenvectors of L_2, where $i, j \in [m]$ and $e_i \otimes e_j$ denotes $\begin{bmatrix} e_i^{(1)} e_j \\ \vdots \\ e_i^{(n)} e_j \end{bmatrix}_{m^2 \times 1}$

Let $e, f \in \{e_i, i \in [n]\}$, we note λ, μ their eigenvalues:

for $1 < k < m$

$$L_2\, e \otimes f = \begin{bmatrix} M & -I & & \\ -I & N & \ddots & \\ & \ddots & N & -I \\ & & -I & M \end{bmatrix} \begin{bmatrix} e^{(1)} f \\ \vdots \\ e^{(n)} f \end{bmatrix} = \begin{bmatrix} M e^{(1)} f - e^{(2)} f \\ -e^{(k-1)} f - e^{(k+1)} f + N e^{(k)} f \\ -e^{(n-1)} + M e^{(n)} f \end{bmatrix}$$

▷ $M e^{(1)} f - e^{(2)} f = e^{(1)} f - e^{(2)} f + L_1 e^{(1)} f = (L_1 e)^{(1)} f + e^{(1)} L_1 f$

(using $M = I + L_1$)

so $M e^{(1)} f - e^{(2)} f = \lambda e^{(1)} f + \mu e^{(1)} f = \underline{\lambda + \mu \quad e^{(1)} f}$

▷ the same way, we can show that:

$- e^{(n-1)} f + M e^{(n)} f = \underline{(\lambda + \mu) \; e^{(n)} f}$

▷ for $k \in]1; n[$,

$- e^{(k-1)} f - e^{(k+1)} f + N e^{(k)} f$

$= (-e^{(k-1)} + 2 e^{(k)} - e^{(k+1)}) f + L_1 e^{(k)} f$

$= (L_1 e)^{(k)} f + e^{(k)} L_1 f$

$= \underline{(\lambda + \mu) \; e^{(k)} f}$

There fore $\underline{L_2 \; e \otimes f = (\lambda + \mu) \; e \otimes f}$ so the $e_i \otimes e_j$ are eigenvectors associated to $(\lambda_i + \lambda_j)$. They are exactly n^2 of them (and they are all different because the eigenvectors of L_1: $e_1, \ldots, e_n$ form a basis of $\mathbb{R}^n$, so the same goes for $(e_i \otimes e_j)_{i,j \in [n]}$

- We have proven that the eigenvalues of L_2 are $\lambda_i + \lambda_j$ for $i,j = 1, \ldots, n$ and that the corresponding eigenvectors are $(e_i \otimes e_j)$

2) First, thanks to the previous question, we can write:

$$L_1 = \sum_{i=1}^{n} \lambda_i \, e_i e_i^T$$

$$L_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} (\lambda_i + \lambda_j)(e_i \otimes e_j)(e_i \otimes e_j)^T$$

{ this can be proven using the fact that the $(e_i)_{i \in [n]}$ and $(e_i \otimes e_j)_{i,j \in [n]}$ are orthonormal basis, and

} that L_1 (resp. L_2) coincides with the sum on the
{ basis $(e_i)_{i\in[n]}$ (resp. $(e_i\otimes e_j)_{i,j\in[n]}$)

and we can also recall that $e_i^T e_j = \delta_{ij}$

$(e_i\otimes e_j)^T(e_k\otimes e_l) = \delta_{ik}\delta_{jl}$

for any $i,j,k,l = 1,\dots,n$

Therefore, for any $m\in\mathbb{N}^*$

$$L_1^m = \sum_{i=1}^{m} \lambda_i^m e_i e_i^T$$

$$L_2^m = \sum_{i=1}^{m}\sum_{j=1}^{m} (\lambda_i+\lambda_j)^m (e_i\otimes e_j)(e_i\otimes e_j)^T$$

and $K_1 = e^{-tL_1} = 1 - tL_1 + \frac{t^2}{2!}L_1^2 + \dots + \frac{t^m}{m!}L_1^m + \dots$

$$= \sum_{i=1}^{m} e^{-t\lambda_i} e_i e_i^T$$

$$K_2 = \sum_{i=1}^{m}\sum_{j=1}^{m} e^{-t(\lambda_i+\lambda_j)} (e_i\otimes e_j)(e_i\otimes e_j)^T$$

Let $(k_1, l_1, k_2, l_2) \in \{1,\dots,n\}$,

$$K_2((k_1,l_1),(k_2,l_2)) = \sum_{i,j} e^{-t(\lambda_i+\lambda_j)} (e_i\otimes e_j)(e_i\otimes e_j)^T_{(k_1,l_1),(k_2,l_2)}$$

$$= \sum_{i,j} e^{-t(\lambda_i+\lambda_j)} e_i^{(k_1)} e_j^{(l_1)} e_i^{(k_2)} e_j^{(l_2)}$$

$$= \left(\sum_{i=1}^{m} e^{-t\lambda_i} e_i^{(k_1)} e_i^{(k_2)}\right)\left(\sum_{j=1}^{m} e^{-t\lambda_j} e_j^{(l_1)} e_j^{(l_2)}\right)$$

$$= K_1(k_1,k_2)\, K_1(l_1,l_2)$$

3) We assume that computing L_1 is done in $O(n^2)$ (so it does not impact the final complexity, since the multiplication takes $O(n^3)$)

L_1 is of size $n \times n$ so computing K_1 is of complexity $O(n^3)$

To compute K_2, using the relation of 2), one only needs n^4 operations, after having computed K_1 (that is done with complexity $O(n^3)$) so the complexity of computing K_2 is $O(n^4)$.

Remarks : - using the symmetry of K_2 we can reduce a bit the number of computations
- using L_2 to compute K_2 would be of complexity $O((n^2)^3) = O(n^6)$, hence the use of K_1 is judicious.

Exercise 1 : B_n - Splines

- First, we deal with the case $\underline{B_1}$.

Take $x_1 = 0$, $x_2 = 1$, $x_3 = 2 \in \mathbb{R}$, and denote K the Gram matrix associated with the kernel $k(x,y) = B_1(x-y)$.

$$K = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{so} \quad K \begin{pmatrix} -1 \\ 2/\sqrt{2} \\ -1 \end{pmatrix} = \begin{pmatrix} -1 + 2/\sqrt{2} \\ -2 + 2/\sqrt{2} \\ -1 + 2/\sqrt{2} \end{pmatrix} = (1-\sqrt{2}) \begin{pmatrix} -1 \\ 2/\sqrt{2} \\ -1 \end{pmatrix}$$

so $\underline{1-\sqrt{2} < 0}$ is an eigenvalue of K so K is not positive semi definite.

Hence the kernel $\underline{k \text{ is not p.d. } (k(x,y) = B_1(x-y))}$

- Now, take $m > 1$ and consider $k(x,y) = B_n(x,y)$, we will show that B_n is continuous.

→ First let us show by induction that $\underline{B_n(x) = 0}$ for any x such that $x < -m$ or $x > m$.

- $B_1(x) = 0$ if $x < -1$ or $x > 1$
- Take $k \geq 1$ and $x \in \mathbb{R}$.

$$B_{k+1}(x) = \int_{-\infty}^{+\infty} B_k(u)\, I(x-u)\, du = \int_{x-1}^{x+1} B_k(u)\, du$$

so for $x < -(k+1)$, $x+1 < -k$ so $B_k(u) = 0$ for any $u \in [x-1, x+1]$ and $B_{k+1}(x) = 0$

the same goes for $x > k+1$.

So $B_{k+1}(x) = 0$ if $x < (k+1)$ or $x > (k+1)$

→ The same way, we can show that $\forall n \geq 1, \forall x \in [-n,n]$ $0 \leq B_n(x) \leq 2^{n-1}$

→ Therefore, $\forall n \in \mathbb{N}$ $\forall p \in \mathbb{N}^*$ B_n is in $L^p(\mathbb{R})$

→ We then use the following theorem:

$\forall p, q \in [1, +\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$, $\forall f \in L^p(\mathbb{R})$, $\forall g \in L^q(\mathbb{R})$

then $f * g$ is uniformly continuous, thus continuous on $\mathbb{R}$

Therefore $\underline{B_n} = \underbrace{B_{n-1}}_{\in L^2(\mathbb{R})} * \underbrace{I}_{\in L^2(\mathbb{R})}$ is continuous on $\mathbb{R}$

- Now we will use Bochner theorem to show when $k(x,y)$ is a p.d. kernel. Let $n \geq 1$ and $K : x, y \mapsto B_n(x-y)$

→ K is a translation invariant kernel and B_n is continuous so the Bochner theorem tells us that

K is p.d $\iff$ B_n is the Fourier-Stieltjes transform of a symmetric and positive finite Borel mesure $\mu \in M(\mathbb{T})$ ($\mathbb{T} = [0, 2\pi]$ with 0 and 2π identified)

→ We denote by $\hat{f}$ the Fourier transform of $f \in L^1(\mathbb{R})$

We have, for any $w \in \mathbb{R}$

$$\hat{I}(w) = \int_{-\infty}^{+\infty} e^{-iwx} I(x)\, dx = \int_{-1}^{1} e^{-iwx}\, dx = \left[\frac{e^{-iwx}}{-iw}\right]_{-1}^{1}$$

$$= \frac{e^{-iw} - e^{iw}}{-iw} = \frac{-2i \sin w}{-iw} = 2\frac{\sin w}{w} = 2 \,\mathrm{sinc}(w)$$

Let $n \geq 1$, $\underline{\hat{B_n}(w)} = \widehat{I^{*n}}(w) = (\hat{I})^n(w) = \underline{2^n(\mathrm{sinc}(w))^n}$

We have that $\mathrm{sinc} \in L^1(\mathbb{R})$ and $|\mathrm{sinc}| \leq 1$ so we also have sinc^m and $2^m \text{sinc}^m \in L^1(\mathbb{R})$

Therefore the inverse Fourier formula holds and

$\forall x \in \mathbb{R}$ $$B_n(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ix\omega}\,\hat{B_n}(\omega)\,d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ix\omega}\, 2^m \,\text{sinc}^m(\omega)\,d\omega$$

change of variable $u = -\omega$

$$= \int_{-\infty}^{+\infty} e^{ix\omega}\left(2^m \text{sinc}^m / 2\pi\right)(\omega)\,d\omega$$

$$= \int_{-\infty}^{+\infty} e^{-ix\omega}\left(2^m \text{sinc}^m / 2\pi\right)(\omega)\,d\omega$$

positive if and only if m is even, because $\text{sinc}(-\pi/2) = -1$

and when m is even $2^m \text{sinc}^m / 2\pi$ is symmetric.

Therefore, using Bochner theorem, K is p.d. $\Leftrightarrow$ m is even

- Then, we know that the RKHS with K as r.k. is

$$\mathcal{H} = \left\{ f \in L^2(\mathbb{R}),\ \|f\|_K^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{|\hat{f}(\omega)|^2}{2^m \text{sinc}^m(\omega)}\,d\omega < \infty \right\}$$

endowed with the inner product

$$\langle f, g \rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\hat{f}(\omega)\,\overline{\hat{g}(\omega)}}{2^m \text{sinc}^m(\omega)}\,d\omega$$

(for any $m \in \mathbb{N}^*$ that is even)